VIROLOGY 212, 741–745 (1995)

SHORT COMMUNICATION

Distant Strains of the Fish Rhabdovirus VHSV Maintain a Sixth Functional Cistron Which Codes for a Nonstructural Protein of Unknown Function

BERNARDO BASURCO and ABDENOUR BENMANSOUR[1]

Unité de Virologie et Immunologie Moléculaires, INRA, 78352 Jouy-En-Josas Cédex, France

Received May 22, 1995; accepted July 24, 1995

We used direct RNA sequencing to determine the genomic organization of the region downstream from the G gene of viral hemorrhagic septicemia virus (VHSV), a fish rhabdovirus. This region contains a gene coding for a protein, identified as nonvirion protein (NV), and the gene coding for the RNA polymerase (L). Thus, VHSV genome organization was confirmed to be 3′-N-P-M-G-NV-L-5′. In both a virulent European (07-71) and an avirulent North American (Makah) strain, the NV gene is transcribed into a small mRNA that codes for a protein of 122 amino acids. It has no significant sequence similarity with the infectious hematopoietic necrosis virus NV protein nor with any other known protein. We expressed the NV protein as a fusion protein with the glutathione *S*-transferase of *Schistosoma japonicum* and used the purified fusion protein to immunize rabbits. The rabbit antiserum precipitated from infected cell extracts — and not from noninfected cells or purified virions — a protein of 14 kDa, well in accordance with the expected NV gene product size. The prediction that the NV protein is a nonstructural protein is supported by its absence from mature virions although it is present in infected cells.

Viral hemorrhagic septicemia virus (VHSV) is a salmonid rhabdovirus that causes severe losses in the European fish farming industry (*1*). Mortality rates among juvenile rainbow trout, *Oncorhynchus mykiss,* can be as high as 90%. VHSV has been isolated in western North America (*2*), and an avirulent strain was obtained from adult coho salmon, *O. kisutch,* during a routine control check (*3*). VHSV belongs to the Rhabdoviridae family. According to its electrophoretic protein pattern, it has been classified in the Lyssavirus genus. Rabies virus (RV), the prototype Lyssavirus, contains a nonsegmented RNA molecule of approximately 12,000 nucleotides (nt) which is sequentially transcribed into a leader RNA and monocistronic mRNAs coding for five structural proteins: the nucleocapsid protein (N), a polymerase-associated protein (P or M1), a matrix protein (M or M2), a glycoprotein (G), and an RNA polymerase (L). RV is also characterized by the presence of an abnormally long (450 nt) noncoding region at the 5′ end of the G gene (*4*). Infectious hematopoietic necrosis virus (INHV), another salmonid fish rhabdovirus, has a functional gene between the G and L regions, which codes for a nonvirion (NV) protein of unknown function (*5*, *6*). The nucleotide sequence determinations of the 3′ half of the genome of VHSV have confirmed a genomic organization similar to the prototype RV and to IHNV (*7–9*). In this present work, we determined the organization of the remaining 5′ half of the genome and found it was similar to IHNV. It is characterized by the presence of a conserved sixth functional cistron at the G–L intergenic region. It codes for a protein which was not detected in mature virions, but was specifically precipitated from infected cell extracts.

We used a synthetic oligonucleotide corresponding to the 3′ coding region of the G mRNA to start direct sequencing on the genomic RNA of VHSV strains originating from Europe (strain 07-71) or the United States (strain Makah). The sequence of approximately 1000 nt downstream from the G gene was determined. This region comprised two open reading frames (ORF). The first one corresponded to the NV gene of IHNV. It had the primary structure characteristics of a VHSV gene (*9*). While the consensus termination sequence AGATAG(A)$_7$ was conserved in both 07-71 and Makah strains, a G for C substitution at the AACA hypothetical transcription start signal was found in the Makah strain (Fig. 1a). The second ORF started at nt 400 and was still continuing after nt 950. It coded for the first 150 amino acids of a protein which had sequence features common to rhabdoviral RNA-dependent RNA polymerase (Fig. 1b). Thus, it most probably represents the 3′ extremity of the L gene.

The deduced amino acid sequences of the NV protein

The nucleotide sequence data reported in this article have been deposited with the GenBank Database under Accession Nos. U28745 and U28746.

[1] To whom correspondence and reprint requests should be addressed. Fax: 33.1.34 65 25 91.

0042-6822/95 $12.00

a

G NV L

07-71: AGATAGAAAAAAA..21..AACATG..363..TGA...18..AGATAGAAAAAAA..32..AACA..61..ATG..430..//..

Makah: AGATAGAAAAAAA..21..AAGATG..363..TGA...18..AGATAGAAAAAAA..//..

b

07-71: MEMFELDR EVH QERLPSECS..... LNSPL NLSLSLQLFGRL.//.

RV: MLDPG..... EVY DDPIDPIELEAEPRGTPTVPNILRNS.....DYN LNSPL IEDPARLMLEWL.//.

VSV: M..... EVH DFETDEFNDFNEDDYATREFLNPDERMTYLNHADYN LNSPL SDDIDNLIRKFNS.//.

FIG. 1. Presence of two genes in the 5′ half of VHSV genome. (a) Genomic organization. Consensus transcription start and termination signals are in bold. Translation start and stop codons are in bold and underlined. Numbers indicate the length in nt between characters. (b) Alignment of the amino-terminal region of the putative L protein of strain 07-71 with RV and VSV RNA polymerases. A stretch of six conserved amino acids (*15*) is indicated in bold. A shorter conserved stretch of three amino acids is underlined. Dots were introduced to optimize the amino acid alignment.

of the 07-71 and Makah strains are shown in Fig. 2. The NV protein is 122 amino acids long and has a calculated molecular weight of 13.6 kDa. The amino acid sequence comparison of the NV protein of the European 07-71 and the American Makah strains showed a 71.5% identity, which was significantly lower than the corresponding percentage identities found for the N, G, P, and M proteins: 91.1, 92.5, 89.7, and 94.5%, respectively (*10, 9*). Forty percent of the substitutions (14 of 35) were localized in the last 26 amino acids, which represented a 54% divergence for this region. In the course of a molecular epidemiology study, we determined the NV gene sequence of 18 different strains isolated in Europe and the United States and representing the four serotypes of VHSV. We confirmed the presence of an intact NV gene in all the analyzed strains as well as a high sequence divergence between the strains from the two continents. The sequences were remarkably conserved, however, among the strains originating from the same continent, regardless of their serotype (data not shown). As also shown in Fig. 2, the NV protein of both VHSV strains had only limited sequence similarities with the putative NV protein of IHNV (GenBank Accession No. X73872). This observation, in addition to the nonconservation of the consensus transcription start signal sequence raised the question of whether this region of the VHSV viral genome is functionally active.

To identify NV mRNA transcripts in VHSV-infected cells, we cloned the region containing the NV gene in both VHSV strains. The cloned cDNA was labeled and used as a probe in Northern blot hybridization. A similarly labeled probe made from the M gene of the 07-71 strain was used as a positive control. The results of the Northern hybridization with the viral RNA or the total RNA prepared from infected and uninfected control cells are shown in Fig. 3. A single band of approximately 0.4 kb was identified with the NV probes in the total RNA from cells infected with both VHSV strains. However, a positive reaction was only observed with the homologous NV probe, while the M probe from strain 07-71 hybridized as well to the total RNA from the Makah strain-infected cells. The negative reaction observed with the NV nonho-

```
nv_07      MATQPGLSTTSF  SPLV LRE MITHRLKFDPSNY  LNC DFD RSD     41
nv_mak     MTTQSAHSTTSF  SPLV LRE MIAYRLTFDPSNY  LNS DLD RSE     41
nv_ihn     mdhrdintnmea----  LRE VLRYKNKVAGHGF  LFD DGD LVW     37

nv_07      ISTV DFFETT LPRILDD LRA STR LPHLHVL DMR ISLL ERTH     82
nv_mak     ISAT DFFETT LSRVLKD LRT STR LPYLHVL DMR ISLL EGTH     82
nv_ihn     REED DATWRR LCDVVNA LIS SKR MQRVLYM DLS ITKG EGhl     78

nv_07      YMFRNVPSSPATTGRLTDPGLV  ISHAEVGLLTR  GSglts        123
nv_mak     YILRNVSSSPATTGRPSDPGLF  ISLEGMKTLTN  GSespp        123
nv_ihn     lfvdlqgtknrlhkeprfrrhl  liedflaypr-------         112
```

FIG. 2. NV protein sequence comparison. Alignment of the deduced amino acid sequences of the NV genes of strains 07-71 and Makah with the NV protein of IHNV (X73872). Optimal alignments were performed with the MACAW 2.0.5 program. Identities are highlighted.

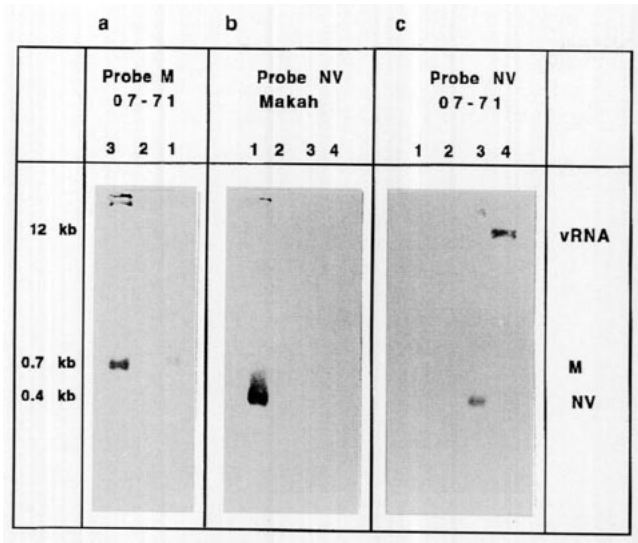

FIG. 3. Identification of NV mRNA by Northern blot analysis. Viral RNA from strain 07-71 (lane 4) and total RNA extracted from EPC cells infected with strain 07-71 (lane 3) or Makah (lane 1) or mock infected EPC cells (lane 2) were hybridized with 07-71 M (a), Makah (b), or 07-71 NV (c) probes.

mologous probes, but not with the M probe, probably resulted from the high stringency of our hybridization conditions, in addition to the lower percentage of sequence identity in the NV gene (76%) versus the M gene (86%). Thus, despite the nonconservation of the transcription start signal, the NV cistron is functional in both strains of VHSV.

No protein band in the range of the expected size of the NV gene product could be detected in Coomassie-stained SDS–PAGE gels of the purified virus. The NV protein could not be detected, either, in [^{35}S]Met-Cys-labeled viral particles, even after a long exposure time (see Fig. 4). To rule out a bias in ^{35}S-labeling due to the small number of methionine and cysteine residues in the NV protein, the labeling of viral particles was also performed with [^{3}H]leucine. Except for the L, G, N, P, and M proteins, no ^{3}H-labeled protein could be detected in the mature virions (data not shown).

The cDNAs containing the coding region of the NV gene of 07-71 and Makah strains were obtained by RT–PCR amplification and were subsequently inserted in frame with the glutathione *S*-transferase (GST) gene coding sequence of a modified pGEX plasmid (Amgen). Positive cultures expressing a shift in the molecular weight of the GST protein were detected by SDS–PAGE and were confirmed to contain GST by Western blotting using an anti-GST serum. The NV moiety of the GST–NV fusion protein was estimated, by deduction, to have a molecular weight of 14 kDa, which is well in agreement with the molecular weight calculated from the sequence. The purification of the GST–NV fusion protein was achieved by affinity chromatography on glutathione–Sepharose 4B. The purified GST–NV fusion protein was used to immunize rabbits. The serum obtained was adsorbed on GST glutathione–Sepharose beads, and was subsequently shown to recognize GST–NV in Western blot, but did not recognize GST alone (data not shown). The GST–NV-specific serum, however, failed to detect any protein in the purified virions or in the infected cell extracts by Western blot analysis. It also failed to label infected cells in an indirect fluorescence test (data not shown). These findings could mean that NV fused to GST was improperly folded and was unable to generate specific antibodies to the native protein. Alternatively, they could also result from an insufficient sensitivity of the Western blot and immunofluorescence tests.

To better investigate the presence of the NV protein in infected cells, specific GST–NV immune and preimmune sera were used in immunoprecipitation tests. A preliminary experiment had shown that a ^{35}S-labeled 14-kDa protein was faintly detected in the precipitate from a 16-hr-p.i. cell extract, but not from 2- to 8-hr-p.i. cell extracts (data not shown). These results were further confirmed in different experiments summarized in Fig. 4. All the major viral proteins, except the L protein, were present in the cell extracts from 12 to 42 hr p.i. (Fig. 4a) and could be precipitated by an anti-VHSV polyclonal serum (Fig. 4c). The L protein was only detected after 12 hr p.i. Similarly, no protein of the expected NV size was precipitated by the GST–NV-specific serum from the 12-hr-p.i. cell extract (Fig. 4c). A protein of approximately 14 kDa was clearly precipitated from the 24-hr-p.i. cell extract by the GST–NV immune serum (Figs. 4d and 4e), but not by the preimmune serum (Fig. 4e). Finally, the GST–NV immune serum failed to precipitate any protein from the ^{35}S-labeled purified virions, confirming that the NV protein was not incorporated into the mature virions (Fig. 4d).

Several faint protein bands were also irregularly detected in the immunoprecipitated complex. Two of them comigrated with the L and N virion proteins (Figs. 4d and 4e) and were more consistantly observed. However, an anti-N monoclonal antibody failed to coprecipitate the NV protein (Fig. 4e), suggesting that these proteins may have been brought down nonspecifically with the complex. To further investigate the binding of viral or cellular proteins to NV, we used the GST–NV fusion protein fixed on glutathione–Sepharose beads to capture any potentially binding protein from infected or noninfected cell extracts. For this purpose, the cell extracts and disrupted virions were first adsorbed on GST–glutathione–Sepharose beads, and then incubated with GST–NV–glutathione–Sepharose or GST–glutathione–Sepharose beads. The captured proteins were resolved by SDS–PAGE and revealed by Western blot analysis with an anti-VHSV polyclonal serum. GST–NV was not able to capture any viral protein (Fig. 5). Surprisingly, GST–NV bound a cellular component of approximately 67 kDa, which was recognized by the anti-VHSV polyclonal serum. This component was also detected in bound material from purified

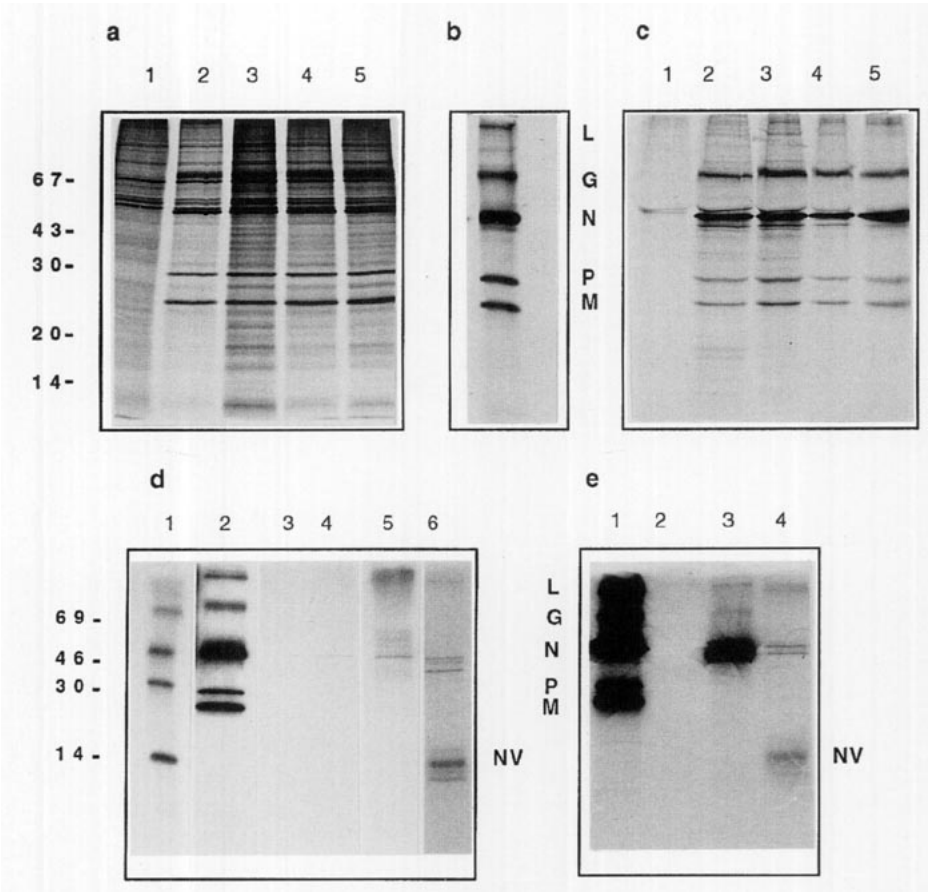

FIG. 4. Identification of the NV protein in VHSV-infected cells. (a) SDS–PAGE of ^{35}S-labeled crude cell extracts. Lane 1, noninfected cells; lane 2, infected cells 12 hr p.i.; lane 3, infected cells 24 hr p.i.; lane 4, infected cells 36 hr p.i.; lane 5, infected cells 42 hr p.i. (b) Purified ^{35}S-labeled virions from supernatant of 42 hr p.i. (c) Lanes 1, 2, 3, 4, 5 same as (a) following immunoprecipitation with an anti-VHSV rabbit polyclonal antibody. (d) Immunoprecipitation with an anti-GST–NV-specific serum. Lane 1, ^{14}C-labeled MW markers, from top to bottom, 69, 46, 30, 14 kDa; lane 2, nonprecipitated ^{35}S-labeled purified virions; lane 3, same as 2 following immunoprecipitation; lane 4, noninfected cell extract; lane 5, 12-hr-p.i. cell extract; lane 6, 24-hr-p.i. cell extract. (e) Lane 1, nonprecipitated ^{35}S-labeled purified virons; lane 2, 24-hr-p.i. cell extract following immunoprecipitation with rabbit preimmune serum; lane 3, 24-hr-p.i. cell extract following immunoprecipitation with an anti-N MAb; lane 4, 24-hr-p.i. cell extract following immunoprecipitation with anti-GST–NV-specific serum.

mature virions, where it was masked by the G protein. It was not metabolically labeled, however, nor it was detected by an anti-G monoclonal antibody (data not shown). Thus, it may represent either a nonprotein component of cellular origin or a protein from the cell culture medium, some of which are known to solidly attach to purified virions. It was not detected on GST alone because the cell extracts were adsorbed on GST–glutathione–Sepharose beads prior to testing.

The availability of the GST–NV fusion protein allowed us to search for the presence of specific antibodies to NV in the sera of trout surviving either experimental or natural infection. A pool of trout sera with a high neutralizing titer was analyzed by Western blotting against the GST–NV fusion protein. No positive reaction was detected (data not shown). This result suggests that the NV protein, probably due to its nonstructural nature, is weakly immunogenic in a natural infection. Alternatively, the NV protein fused to GST could be nonantigenic for trout antibodies.

VHSV and IHNV are strictly dependent upon a low (14°) temperature for efficient replication *in vivo* and *in vitro.* Kurath *et al.* (*5*) suggested that the sixth additional gene of IHNV, which is not present in mammalian rhabdoviruses, might be implicated in this low-temperature dependence. The availability in our laboratory of a temperature-stable variant of strain 07-71 of VHSV (*11*), which was able to replicate efficiently at 25°, allowed us to test this hypothesis. The NV gene of the temperature variant was amplified by RT–PCR, and the amplified cDNA product was sequenced. Compared to the parental 07-71 sequence, no variation could be detected in the NV protein of the temperature variant (data not shown). Thus, higher temperature adaptation had no influence on the NV sequence. Consequently, the NV protein is probably not directly implicated in the temperature dependence of VHSV.

We have located, by direct RNA sequencing, two genes in the 5′ region downstream from the G gene of VHSV. The first one was identified as a sixth additional

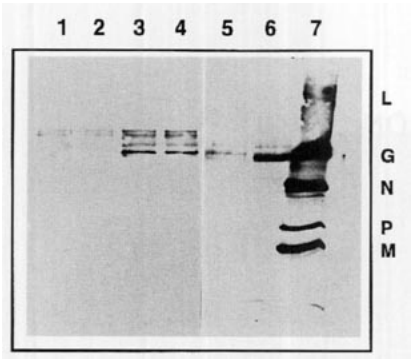

FIG. 5. GST–NV binding. Noninfected cell extract (lanes 1 and 3), 24-hr infected cell extract (lanes 2 and 4), or disrupted purified virion (lanes 5, 6, and 7) were first adsorbed on GST–glutathione–Sepharose beads, then incubated with GST–glutathione–Sepharose beads (lanes 1, 2, and 5) or GST–NV–gluthatione–Sepharose beads (lanes 3, 4, and 6). After thorough washing, the captured proteins were disrupted in Laemmli's buffer, resolved on SDS–PAGE, electroblotted on nitrocellulose, and immunodetected with an anti-VHSV rabbit polyclonal serum according to the ECL Western blotting protocol (Amersham).

gene coding for a nonvirion protein, and the second was identified as the RNA-directed RNA polymerase gene. Thus, the VHSV genome organization was confirmed to be 3′-N-P-M-G-NV-L-5′. In reviewing the homology of VHSV with other rhabdoviruses, it was found that IHNV was the most closely related virus. Both are fish rhabdoviruses which cause high mortality levels in salmonid species and follow a similar pathogenic course (*12*). The presence of a sixth gene in the same position (*5*) provided an additional analogy between the two salmonid rhabdoviruses. In addition to a similar genome organization, VHSV and IHNV conserved the same transcription start and termination signals (*9, 13*) and have significant percentages of identity between most of their structural proteins (N, 39%; P, 37%; M, 37%; G, 39%). Surprisingly, there is only limited sequence similarity between NV proteins of VHSV and IHNV, indicating that the structure, rather than the sequence, is important for the NV protein function. A protein with the expected NV gene product size was precipitated from VHSV-infected cell extracts. However, it was only detected through the most sensitive techniques, indicating that the NV protein is probably needed in catalytic amounts. Alternatively, the NV protein may be nonfunctional, its low level of accumulation, late in the viral cycle, being of trivial consequence in conditions of high viral outputs. It is known from RV reverse genetics that the corresponding remnant gene can be either completely deleted or turned on to be functional without influencing RV biology (*14*). Supporting this hypothesis, the NV protein could not be localized in infected cells, and no special pattern of association with viral or cellular proteins was demonstrated. On the other hand, conservation of a functional gene in geographically distant strains of VHSV and in IHNV and the striking sequence stability among isolates of the same lineage are strong indications of the biological importance of the NV gene function; however, its nature remains to be elucidated.

ACKNOWLEDGMENTS

This work was financially supported by the Institut National de la Recherche Agronomique, France. Dr. B. Basurco was a recipient of a postdoctoral fellowship from the Instituto Nacional de Investigaciones Agrarias, Spain. The excellent technical assistance of Anne F. Monnier and Patrice Vende is gratefully acknowledged.

REFERENCES

1. de Kinkelin, P., Chilmonczyck, S., Dorson, M., Le Berre, M., and Baudouy, A. M., *In* "Munich Symposia on Microbiology: Mechanisms of Viral Pathogenesis and Virulence" (P. A. Bachmann, Ed.), pp. 357–375. WHO Collaborating Center, Munich, 1979.
2. Hopper, K., *Am. Fish. Soc. Fish Health Sect. Newsl.* **17,** 1–2 (1989).
3. Brunson, R., True, K., and Yancey, J., *Am. Fish. Soc. Fish Health Sect. Newsl.* **17,** 3–4 (1989).
4. Tordo, N., Poch, O., Ermine, A., Keith, G., and Rougeon, F., *Proc. Natl. Acad. Sci. USA* **83,** 3914–3918 (1986).
5. Kurath, G., Ahern, K. D., Pearson, G. D., and Leong, J. C., *J. Virol.* **53,** 469–476 (1985).
6. Kurath, G., and Leong, J. C., *J. Virol.* **53,** 462–468 (1985).
7. Bernard, J., Lecoq-Xhonneux, F., Rossius, M., Thiry, M. E., and De Kinkelin, P., *J. Gen. Virol.* **71,** 1669–1674 (1990).
8. Thirty, M., Lecoq-Xhonneux, F., Dheur, I., Renard, A., and de Kinkelin, P., *Biochim. Biophys. Acta* **1090,** 345–347 (1991).
9. Benmansour, A., Paubert, G., Bernard, J., and de Kinkelin, P., *Virology* **198,** 602–612 (1994).
10. Bernard, J., Bremont, M., and Winton, J., *J. Gen. Virol.* **73,** 1011–1014 (1992).
11. de Kinkelin, P., *In* "Fish Vaccination" (A. E. Ellis, Ed.), pp. 172–192. Academic Press, London, 1988.
12. Wolf, K., "Fish viruses and fish viral diseases." Cornell Univ. Press, New York, 1988.
13. Koener, J. F., Passavant, C. W., Kurath, G., and Leong, J. C., *J. Virol.* **61,** 1342–1349 (1987).
14. Schnell, M. J., Mebatsion, T., and Conzelman, K. K., *EMBO J.* **13,** 4195–4205 (1994).
15. Poch, O., Blumberg, B. M., Bougueleret, L., and Tordo, N., *J. Gen. Virol.* **71,** 1153–1162 (1990).